Taming the complexity of biochemical networks through model reduction and tropical geometry

Ovidiu Radulescu

Université de Montpellier, DIMNP UMR 5235

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = P_i(\boldsymbol{x}) = \sum_{j=1}^{M_i} k_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}}, \; \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} = x_1^{\alpha_{ij}^1} \ldots x_n^{\alpha_{ij}^n}.$$

$\boldsymbol{\alpha}_{ij} \in \mathbb{N}^n$ are multi-indices, $k_{ij} \in \mathbb{R}$ are kinetic parameters.

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = P_i(\boldsymbol{x}) = \sum_{j=1}^{M_i} k_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}}, \ \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} = x_1^{\alpha_{ij}^1} \dots x_n^{\alpha_{ij}^n}.$$

$\boldsymbol{\alpha}_{ij} \in \mathbb{N}^n$ are multi-indices, $k_{ij} \in \mathbb{R}$ are kinetic parameters.

› How to identify fast/slow variables? Problem : timescales are not just inverses of k_{ij}.

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = P_i(\boldsymbol{x}) = \sum_{j=1}^{M_i} k_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}}, \ \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} = x_1^{\alpha_{ij}^1} \ldots x_n^{\alpha_{ij}^n}.$$

$\boldsymbol{\alpha}_{ij} \in \mathbb{N}^n$ are multi-indices, $k_{ij} \in \mathbb{R}$ are kinetic parameters.

› How to identify fast/slow variables? Problem : timescales are not just inverses of k_{ij}.

› Which details of the model can be neglected? Problem : small terms can not always be neglected.

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = P_i(\boldsymbol{x}) = \sum_{j=1}^{M_i} k_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}}, \ \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} = x_1^{\alpha_{ij}^1} \dots x_n^{\alpha_{ij}^n}.$$

$\boldsymbol{\alpha}_{ij} \in \mathbb{N}^n$ are multi-indices, $k_{ij} \in \mathbb{R}$ are kinetic parameters.

› How to identify fast/slow variables? Problem : timescales are not just inverses of k_{ij}.

› Which details of the model can be neglected? Problem : small terms can not always be neglected.

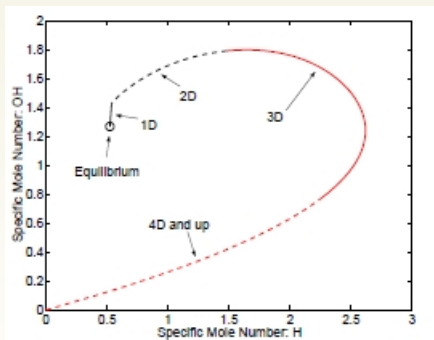› Can one provide symbolic descriptions of the reduced model dynamics?
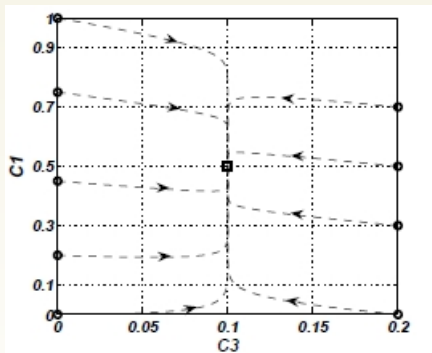
from Chiavazzo et al Comm.Comp.Phys. 2007

from Hung and Sheperd 12th Ann. Int. Detonation Symp. 2002

In fast/slow systems, fast variables relax quickly, then they are slaved : the system moves on a low dimensional invariant manifold describing the reduced model. The reduced model can change with time.

$$\frac{dx}{dt} = \frac{1}{\varepsilon} f(x,y) \text{(1: fast)}$$

$$\frac{dy}{dt} = g(x,y) \text{(2: slow)}$$

Tikhonov : *If for any y the dynamics (1) has a hyperbolic point attractor, then after a fast transition the system evolves according to:*

$$\frac{dy}{dt} = g(x,y)$$

and

f(x,y)=0 the fast variables are slaved by slow ones

Several steps:

Several steps:

› Determine the fast/slow decomposition : partially solved by CSP (Lam and Goussis 1994), implemented in COPASI.

Several steps:

› Determine the fast/slow decomposition : partially solved by CSP (Lam and Goussis 1994), implemented in COPASI.

› Solve $f(x, y) = 0$ for x : hard, few symbolic methods(Grobner bases, sparse polynomial systems (Grigoriev, Weber 2012)), no implementation.

Several steps:

› Determine the fast/slow decomposition : partially solved by CSP (Lam and Goussis 1994), implemented in COPASI.

› Solve $f(x, y) = 0$ for x : hard, few symbolic methods(Grobner bases, sparse polynomial systems (Grigoriev, Weber 2012)), no implementation.

› Pool reactions, find effective rate laws in y : use elementary modes (Radulescu et al 2008).

Several steps:

› Determine the fast/slow decomposition : partially solved by CSP (Lam and Goussis 1994), implemented in COPASI.

› Solve $f(x, y) = 0$ for x : hard, few symbolic methods(Grobner bases, sparse polynomial systems (Grigoriev, Weber 2012)), no implementation.

› Pool reactions, find effective rate laws in y : use elementary modes (Radulescu et al 2008).

› Pool species : use conservation laws (Gorban, Radulescu, Zinovyev 2010).

Fast, low concentration radicals
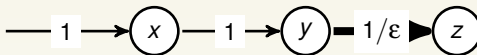
$$\frac{dx}{dt} = 1 - x$$

$$\frac{dy}{dt} = x - \frac{1}{\varepsilon}y$$

$$\frac{dz}{dt} = \frac{1}{\varepsilon}y - z$$

$$y = \varepsilon x$$

y is fast, *x*, *z* are slow.
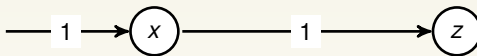
$$\frac{dx}{dt} = 1 - x$$

$$\frac{dz}{dt} = x - z$$

Reduced model : prune fast species, replace fast sub-system by elementary modes
(reaction pooling)
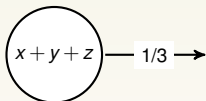
Fast cycles (all the species in the cycle are fast)
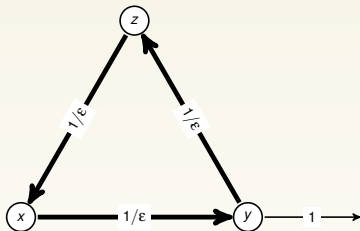
$$\frac{dx}{dt} = \frac{1}{\varepsilon}(z - x)$$

$$\frac{dy}{dt} = \frac{1}{\varepsilon}(x - y) - y$$

$$\frac{dz}{dt} = \frac{1}{\varepsilon}(y - z)$$

Conservation laws of the fast subsystem that are not conserved by the full system are slow.

$$\left(x+y+z\right) - 1/3 \longrightarrow$$

$$\frac{d(x+y+z)}{dt} = -1/3(x+y+z)$$

Reduced model : prune fast reactions, pool species.

Let x^* be the solution of $f(x, y(t)) = 0$ where $y(t)$ are simulated trajectories. The slowness index for x is $I(t) = |\log(x(t)/x^*(t))|$.
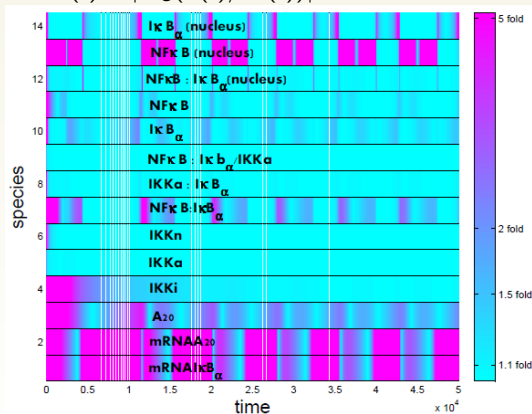
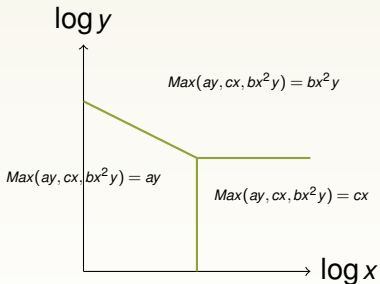Slowness index calculated for the species in the NFκB model BIOMD0000000226.
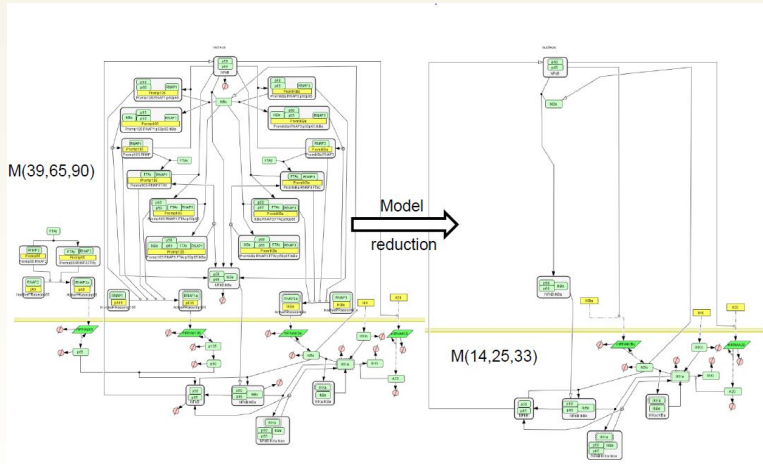
Reduction of NF-κB signaling (BIOMD0000000227 in Biomodels database)

Consider a polynomial in two variables, say $ay + cx + bx^2y$.

Max-plus algebra $+ \rightarrow Max$, $\times \rightarrow +$.

A tropical curve is the set of points where the corresponding max-plus polynomial $Max(log(a) + y, log(c) + x, log(b) + 2log(x) + log(y))$ is not smooth.

A tropical curve is an algebraic curve defined over the max-plus semifield.

› Solve $\varepsilon y - x + \varepsilon^2 x^2 y = 0$, ε small positive parameter

- › Solve $\varepsilon y - x + \varepsilon^2 x^2 y = 0$, ε small positive parameter

- › Kapranov theorem : there are solutions $x = x_1 \varepsilon^{a_1} + x_2 \varepsilon^{a_2} + \ldots$,
  $y = y_1 \varepsilon^{b_1} + y_2 \varepsilon^{b_2} + \ldots$, where $a_1, a_2, \ldots, b_1, b_2, \ldots$ are increasing rational numbers with a common denominator.

› Solve $\varepsilon y - x + \varepsilon^2 x^2 y = 0$, ε small positive parameter

› Kapranov theorem : there are solutions $x = x_1 \varepsilon^{a_1} + x_2 \varepsilon^{a_2} + \ldots$,
  $y = y_1 \varepsilon^{b_1} + y_2 \varepsilon^{b_2} + \ldots$, where $a_1, a_2, \ldots, b_1, b_2, \ldots$ are increasing rational
  numbers with a common denominator.

› must have, at lowest order $y_1 \varepsilon^{1+b_1} - x_1 \varepsilon^{a_1} + x_1^2 y_1 \varepsilon^{2a_1+b_1} = 0$

- › Solve $\varepsilon y - x + \varepsilon^2 x^2 y = 0$, ε small positive parameter

- › Kapranov theorem : there are solutions $x = x_1 \varepsilon^{a_1} + x_2 \varepsilon^{a_2} + \ldots$,
  $y = y_1 \varepsilon^{b_1} + y_2 \varepsilon^{b_2} + \ldots$, where $a_1, a_2, \ldots, b_1, b_2, \ldots$ are increasing rational numbers with a common denominator.

- › must have, at lowest order $y_1 \varepsilon^{1+b_1} - x_1 \varepsilon^{a_1} + x_1^2 y_1 \varepsilon^{2a_1+b_1} = 0$

- › possible, if $min(1 + b_1, a_1, 2a_1 + b_1)$ is attained at least twice, which defines the tropical curve

› Solve $\varepsilon y - x + \varepsilon^2 x^2 y = 0$, ε small positive parameter

› Kapranov theorem : there are solutions $x = x_1 \varepsilon^{a_1} + x_2 \varepsilon^{a_2} + \ldots$, $y = y_1 \varepsilon^{b_1} + y_2 \varepsilon^{b_2} + \ldots$, where $a_1, a_2, \ldots, b_1, b_2, \ldots$ are increasing rational numbers with a common denominator.

› must have, at lowest order $y_1 \varepsilon^{1+b_1} - x_1 \varepsilon^{a_1} + x_1^2 y_1 \varepsilon^{2a_1+b_1} = 0$

› possible, if $min(1 + b_1, a_1, 2a_1 + b_1)$ is attained at least twice, which defines the tropical curve

› to have real solutions need $a_1 = min(1 + b_1, 2a_1 + b_1)$

Tropical equilibration: the positive and negative dominant terms have the same orders.

› For systems, tropical equilibration condition is necessary but not sufficient.

› For systems, tropical equilibration condition is necessary but not sufficient.

›

$$y - x - \varepsilon x^4 = 0$$
$$x - y + \varepsilon y^2 = 0$$
$$x = \varepsilon^{a_1}, y = \varepsilon^{a_2}$$

› For systems, tropical equilibration condition is necessary but not sufficient.

›

$$y - x - \varepsilon x^4 = 0$$
$$x - y + \varepsilon y^2 = 0$$
$$x = \varepsilon^{a_1}, y = \varepsilon^{a_2}$$

› Infinite branch, $a_1 = a_2 \geq -1/3$ and isolated solution $a_1 = -1/2$, $a_2 = -1$

› For systems, tropical equilibration condition is necessary but not sufficient.

›

$$y - x - \varepsilon x^4 = 0$$
$$x - y + \varepsilon y^2 = 0$$
$$x = \varepsilon^{a_1}, y = \varepsilon^{a_2}$$

› Infinite branch, $a_1 = a_2 \geq -1/3$ and isolated solution $a_1 = -1/2$, $a_2 = -1$

› Only $(0,0)$ and $(-1/2, -1)$ can be lifted to Newton-Puiseux series.

› For systems, tropical equilibration condition is necessary but not sufficient.

›

$$y - x - \varepsilon x^4 = 0$$
$$x - y + \varepsilon y^2 = 0$$
$$x = \varepsilon^{a_1}, y = \varepsilon^{a_2}$$

› Infinite branch, $a_1 = a_2 \geq -1/3$ and isolated solution $a_1 = -1/2$, $a_2 = -1$

› Only $(0,0)$ and $(-1/2, -1)$ can be lifted to Newton-Puiseux series.

› By adding the equations, get $y^2 - x^4 = 0$ leading to $a_1 = a_2/2$. Only two solutions left, $(0,0)$ and $(-1/2, -1)$.

› For systems, tropical equilibration condition is necessary but not sufficient.

›

$$y - x - \varepsilon x^4 = 0$$
$$x - y + \varepsilon y^2 = 0$$
$$x = \varepsilon^{a_1}, y = \varepsilon^{a_2}$$

› Infinite branch, $a_1 = a_2 \geq -1/3$ and isolated solution $a_1 = -1/2$, $a_2 = -1$

› Only $(0,0)$ and $(-1/2,-1)$ can be lifted to Newton-Puiseux series.

› By adding the equations, get $y^2 - x^4 = 0$ leading to $a_1 = a_2/2$. Only two solutions left, $(0,0)$ and $(-1/2,-1)$.

› In general, have to compute a tropical basis : hard.

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = \sum_{j=1}^{M_i} k_{ij} x^{\boldsymbol{\alpha}_{ij}}$$

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = \sum_{j=1}^{M_i} k_{ij} x^{\alpha_{ij}}$$

pick $\varepsilon > 0$ is a small positive number.

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = \sum_{j=1}^{M_i} k_{ij} x^{\boldsymbol{\alpha}_{ij}}$$

pick $\varepsilon > 0$ is a small positive number.

$$\gamma_{ij} = round(\log(k_{ij})/\log(\varepsilon))$$

$$k_{ij} = \bar{k}_{ij} \varepsilon^{\gamma_{ij}}$$

Polynomial kinetics (e.g. mass action law)

$$\frac{dx_i}{dt} = \sum_{j=1}^{M_i} k_{ij} x^{\alpha_{ij}}$$

pick $\varepsilon > 0$ is a small positive number.

$$\gamma_{ij} = round(\log(k_{ij})/\log(\varepsilon))$$

$$k_{ij} = \bar{k}_{ij} \varepsilon^{\gamma_{ij}}$$

Define $x_i = \varepsilon^{a_i} \bar{x}_i$ where a_i are unknown.

$$\frac{d\bar{x}_i}{dt} = \sum_{j=1}^{M_i} \varepsilon^{\mu_{ij}} \bar{k}_{ij} \bar{\mathbf{x}}^{\alpha^{ij}},$$

$$\mu_{ij} = \gamma_{ij} + \sum_{l=1}^{n} \alpha_l^{ij} a_l - a_i.$$

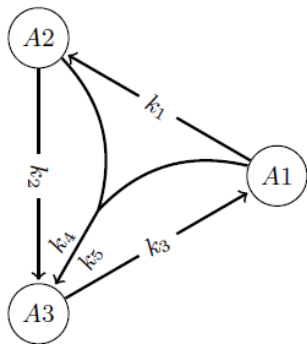At least two terms, of opposite signs have the same degree in ε.

Find a pair (j, j'), $j \neq j'$ such that

i) $\mu_{ij} = \mu_{ij'}$,

ii) $\mu_{ij} \leq \mu_{il}$ for all $l \neq j, j'$,

iii) $k_{ij} k_{ij'} < 0$.

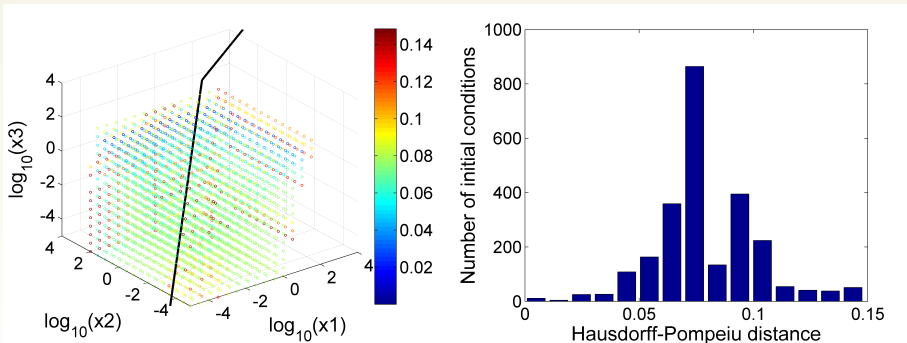The minimum degree is attained at least twice and the minimal degree terms have opposite signs.

$$\frac{d\bar{x}_i}{dt} = \varepsilon^{\mu_i} (|\bar{k}_{ij}| \bar{\mathbf{x}}^{\alpha^{ij}} - |\bar{k}_{ij'}| \bar{\mathbf{x}}^{\alpha^{ij'}})$$
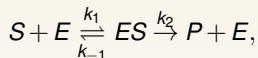
The tropically truncated system is generically binomial (toric system).
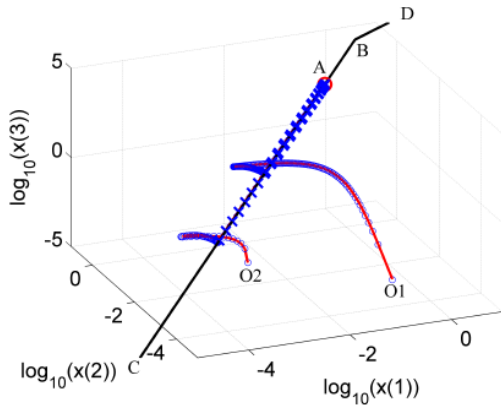
$\gamma_1 = 0, \gamma_2 = 1, \gamma_3 = \gamma_4 = 2, \gamma_5 = 3$

$\min(\gamma_1 + a_1, \gamma_4 + a_1 + a_2) = \min(\gamma_3, \gamma_5) + a_3 = \min(\gamma_2 + a_2, \gamma_4 + a_1 + a_2) = \min(\gamma_1 + a_1, \gamma_5 + a_3)$

Michaelis-Menten enzymatic reaction

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} P + E,$$

Using the conserved quantities $e_0 = [E] + [ES]$, $s_0 = [S] + [ES] + [P]$, we get the system

$$x' = -k_1 x (e_0 - y) + k_{-1} y, \quad y' = k_1 x (e_0 - y) - (k_{-1} + k_2) y.$$

where $x = [S]$ and $y = [SE]$,
and some constraints

$$0 \leq y \leq e_0, 0 \leq x + y \leq s_0, 0 \leq x.$$

After rescalings $x = \bar{x}\varepsilon^{a_1}$, $y = \bar{y}\varepsilon^{a_2}$, $k_1 = \bar{k}_1\varepsilon^{\gamma_1}$, $k_{-1} = \bar{k}_{-1}\varepsilon^{\gamma_{-1}}$, $e_0 = \bar{e}_0\varepsilon^{\gamma_e}$, $s_0 = \bar{s}_0\varepsilon^{\gamma_s}$, we get
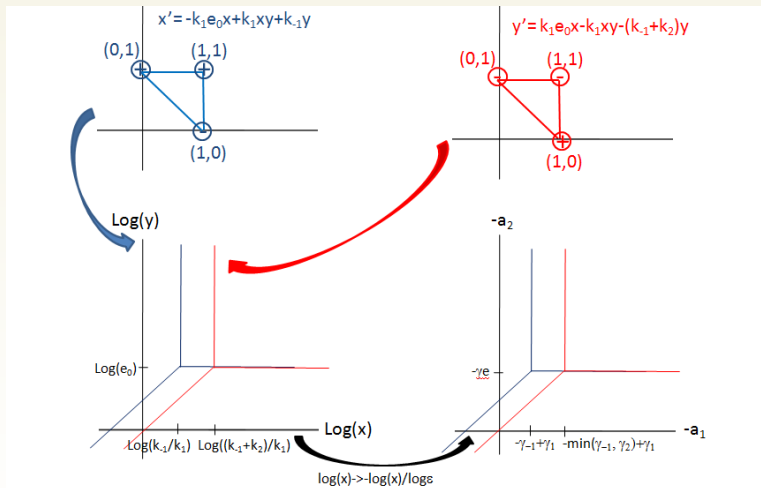
$$
\begin{aligned}
\bar{x}' &= -\bar{k}_1\bar{e}_0\varepsilon^{\gamma_1+\gamma_e}\bar{x} + \bar{k}_1\varepsilon^{\gamma_1+a_2}\bar{x}\bar{y} + \bar{k}_{-1}\varepsilon^{\gamma_{-1}+a_2-a_1}\bar{y}, \\
\bar{y}' &= \bar{k}_1\bar{e}_0\varepsilon^{\gamma_1+\gamma_e+a_1-a_2}\bar{x} - \bar{k}_1\varepsilon^{\gamma_1+a_1}\bar{x}\bar{y} - (\bar{k}_{-1}\varepsilon^{\gamma_{-1}} + \bar{k}_2\varepsilon^{\gamma_2})\bar{y}.
\end{aligned}
$$

that leads to the equilibration equations

$$
\begin{aligned}
\gamma_1 + \gamma_e &= \min(\gamma_1 + a_2, \gamma_{-1} + a_2 - a_1), \\
\gamma_1 + \gamma_e + a_1 - a_2 &= \min(\gamma_1 + a_1, \min(\gamma_{-1}, \gamma_2)).
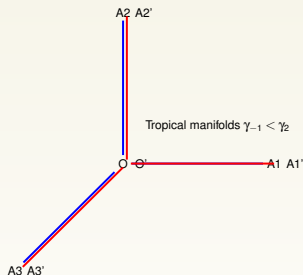\end{aligned}
$$

Tropical manifolds $\gamma_{-1} < \gamma_2$

No	Condition	Truncated system	Regime
1	$a_1 < \gamma_{-1} - \gamma_1$ $a_2 = \gamma_e$	$x' = \varepsilon^{\gamma_1 + \gamma_e} \left(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_1 \bar{x} \bar{y} \right)$ $y' = \varepsilon^{\gamma_1 + a_1} \left(\bar{k}_1 \bar{e}_0 \bar{x} - \bar{k}_1 \bar{x} \bar{y} \right)$	QE saturated
2	$a_1 > \gamma_{-1} - \gamma_1$ $a_2 = a_1 + \gamma_e + \gamma_1 - \gamma_{-1}$	$x' = \varepsilon^{\gamma_1 + \gamma_e} \left(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_{-1} \bar{y} \right)$ $y' = -\varepsilon^{\gamma - 1} \left(\bar{k}_1 \bar{e}_0 \bar{x} - \bar{k}_{-1} \bar{y} \right)$	QE linear

Tropical manifolds $\gamma_{-1} > \gamma_2$

No	Condition	Truncated system	Regime
1	$a_1 < \gamma_2 - \gamma_1$ $a_2 = \gamma_e$	$x' = \varepsilon^{\gamma_1 + \gamma_e}(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_1 \bar{x}\bar{y})$ $y' = \varepsilon^{\gamma_1 + a_1}(\bar{k}_1 \bar{e}_0 \bar{x} - \bar{k}_1 \bar{x}\bar{y})$	QE
2	$\gamma_2 - \gamma_1 < a_1 < \gamma_{-1} - \gamma_1$ $a_2 = \gamma_e$	$x' = \varepsilon^{\gamma_1 + \gamma_e}(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_1 \bar{x}\bar{y})$ $y' = -\varepsilon^{\gamma_2} \bar{k}_2 \bar{y}$	x QSS
3	$a_1 > \gamma_{-1} - \gamma_1$ $a_2 = a_1 + \gamma_e + \gamma_1 - \gamma_{-1}$	$x' = \varepsilon^{\gamma_1 + \gamma_e}(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_{-1} \bar{y})$ $y' = -\varepsilon^{\gamma_2} \bar{k}_2 \bar{y}$	x QSS
4	$a_1 > \gamma_2 - \gamma_1$ $a_2 = a_1 + \gamma_e + \gamma_1 - \gamma_2$	$x' = -\varepsilon^{\gamma_1 + \gamma_e} \bar{k}_1 \bar{e}_0 \bar{x}$ $y' = \varepsilon^{\gamma_2}(\bar{k}_1 \bar{e}_0 \bar{x} - \bar{k}_2 \bar{y})$	y QSS

TTS at quasi-equilibrium : pruning
$\bar{x}' = \varepsilon^{\gamma_1 + \gamma_e}(-\bar{k}_1 \bar{e}_0 \bar{x} + \bar{k}_{-1} \bar{y})$
$\bar{y}' = \varepsilon^{\gamma_{-1}}(\bar{k}_1 \bar{e}_0 \bar{x} - \bar{k}_{-1} \bar{y})$

New slow variable $z = x + y$: pooling
$\bar{z}' = -\varepsilon^{\gamma_2 + a_2 - \gamma_s} \bar{k}_2 \bar{y}$.

Eliminate x, y, obtain the reduced model
$z' = -k_2/(1 + k_{-1}/(k_1 e_0))z$.
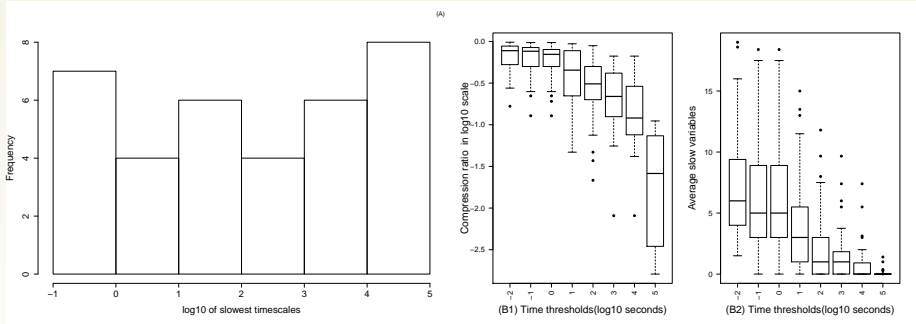
436 curated models, 55 have purely polynomial kinetics.

Two methods :

Reified constraints

Found	# models	Variables (avg/min/max)	Time (avg/min/max)
yes	23	17.348/3/ 86	0.486/0.004/2.803
no	32	17.812/1/194	0.099/0.000/1.934

Newton polytope

ε value	Total models considered	Timed-out models	Models without tropical solutions	Models with tropical solutions	Average running time (in secs)
1/5	53	16	0	37	405.50
1/7	53	16	0	37	734.74
1/9	53	16	0	37	716.40
1/23	50	17	1	32	889.59

(A)

$\mu_{\text{threshold}} = -\log(\theta/\tau)/\log(\varepsilon)$, τ time units.
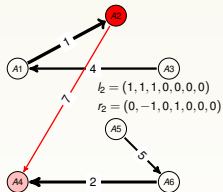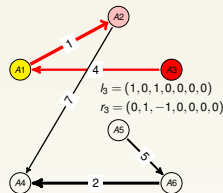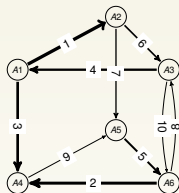
Species with $\mu_i - a_i \geq \mu_{\text{threshold}}$ are slow.

At timescales of $1000s$ models have median numbers of 2 slow variables.
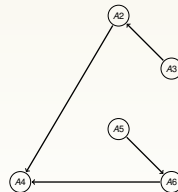
Any monomolecular network with separated constants can be reduced to an acyclic, deterministic digraph (Gorban and Radulescu 2007). The symbolic dynamics can described by a deterministic finite state automaton. The number of states is at most the number of species.

› Main idea : use tropical equilibration branches (slow) as proxys for metastability.

› Main idea : use tropical equilibration branches (slow) as proxys for metastability.

› Branches are defined as all tropical equilibration leading to the same fast truncated system; computed by post-processing of algorithm output.

› Main idea : use tropical equilibration branches (slow) as proxys for metastability.

› Branches are defined as all tropical equilibration leading to the
<u>same fast truncated system</u>; computed by post-processing of algorithm output.

› We define minimal branches : minimal truncated system.
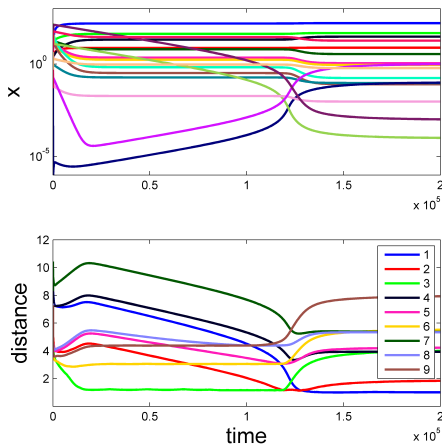
› Main idea : use tropical equilibration branches (slow) as proxys for metastability.

› Branches are defined as all tropical equilibration leading to the
<u>same fast truncated system</u>; computed by post-processing of algorithm output.

› We define minimal branches : minimal truncated system.

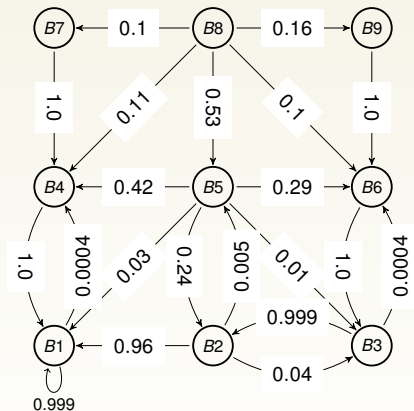› Minimal branches are algebraically connected if they have a common overset.
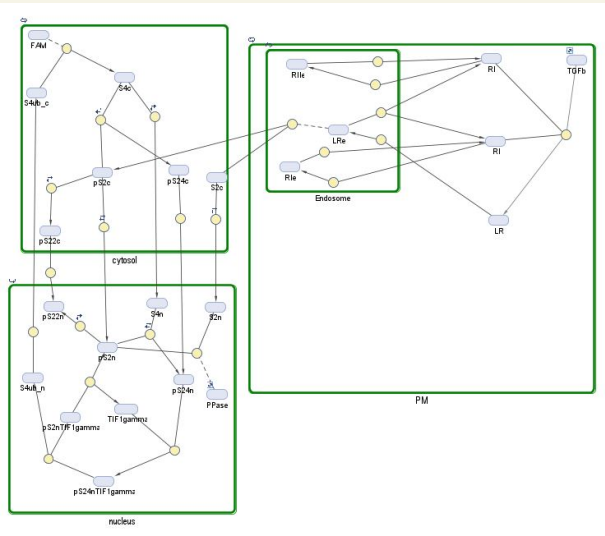
TGFb signaling model.
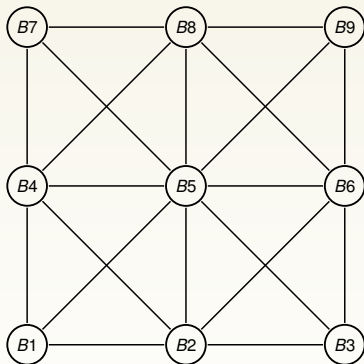
Minimal tropical branches and algebraic connections.

Numerically computed trajectories and distances to minimal branches.

Stochastic finite state automaton learned from trajectories with various initial conditions.

› Solving the tropical equilibration problem allows fast/slow decomposition and model reduction of biochemical reactions networks.

› We have two methods for solving the tropical equilibration problem, a first one by reformulating it as a constraint satisfaction problem and a second one based on the Newton polyhedron.

› Tropical methods can be also used to coarse grain the dynamics of a smooth biochemical reaction network to a discrete symbolic dynamics of a finite state automaton.
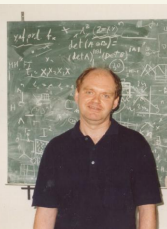
Alexander Gorban   Dima Grigoriev   Vincent Noel   Sergei Vakulenko   Andrei Zinovyev

Francois Fages   Sylvain Soliman   Satya Swarup Samal